Resolution enhancement of lung 4D-CT data using multiscale interphase iterative nonlocal means

Yu Zhang
School of Biomedical Engineering, Southern Medical University, Guangzhou 510515, China and Department of Radiology and BRIC, University of North Carolina, Chapel Hill, North Carolina 27599

Pew-Thian Yap and Guorong Wu
Department of Radiology and BRIC, University of North Carolina, Chapel Hill, North Carolina 27599

Qianjin Feng
School of Biomedical Engineering, Southern Medical University, Guangzhou 510515, China

Jun Lian
Department of Radiation Oncology, University of North Carolina, Chapel Hill, North Carolina 27599

Wufan Chen
School of Biomedical Engineering, Southern Medical University, Guangzhou 510515, China

Dinggang Shen
Department of Radiology and BRIC, University of North Carolina, Chapel Hill, North Carolina 27599 and Department of Brain and Cognitive Engineering, Korea University, Seoul 136-713, Korea

(Received 1 January 2013; revised 18 February 2013; accepted for publication 8 April 2013; published 1 May 2013)

Purpose: Four-dimensional computer tomography (4D-CT) has been widely used in lung cancer radiotherapy due to its capability in providing important tumor motion information. However, the prolonged scanning duration required by 4D-CT causes considerable increase in radiation dose. To minimize the radiation-related health risk, radiation dose is often reduced at the expense of interslice spatial resolution. However, inadequate resolution in 4D-CT causes artifacts and increases uncertainty in tumor localization, which eventually results in extra damages of healthy tissues during radiotherapy. In this paper, the authors propose a novel postprocessing algorithm to enhance the resolution of lung 4D-CT data.

Methods: The authors’ premise is that anatomical information missing in one phase can be recovered from the complementary information embedded in other phases. The authors employ a patch-based mechanism to propagate information across phases for the reconstruction of intermediate slices in the longitudinal direction, where resolution is normally the lowest. Specifically, the structurally matching and spatially nearby patches are combined for reconstruction of each patch. For greater sensitivity to anatomical details, the authors employ a quad-tree technique to adaptively partition the image for more fine-grained refinement. The authors further devise an iterative strategy for significant enhancement of anatomical details.

Results: The authors evaluated their algorithm using a publicly available lung data that consist of 10 4D-CT cases. The authors’ algorithm gives very promising results with significantly enhanced image structures and much less artifacts. Quantitative analysis shows that the authors’ algorithm increases peak signal-to-noise ratio by 3–4 dB and the structural similarity index by 3%–5% when compared with the standard interpolation-based algorithms.

Conclusions: The authors have developed a new algorithm to improve the resolution of 4D-CT. It outperforms the conventional interpolation-based approaches by producing images with the markedly improved structural clarity and greatly reduced artifacts. © 2013 American Association of Physicists in Medicine. [http://dx.doi.org/10.1118/1.4802747]

Key words: lung 4D-CT, resolution enhancement, nonlocal means

I. INTRODUCTION

Four-dimensional computer tomography (4D-CT) is becoming increasingly popular in lung cancer treatment for providing respiratory-related information that is essential for guiding radiation therapy. 4D-CT is usually obtained by sorting multiple free-breathing CT segments in relation to couch position and tidal volume. Due to the risk of radiation, only a limited number of CT segments are usually acquired, which often results in low resolution (LR) along the interslice direction. This LR data are usually plagued with visible imaging artifacts such as vessel discontinuity and partial volume. More importantly, insufficient resolution further distorts the shape of a tumor and may eventually interfere with the target definition of radiation therapy planning. In recent years, several free-breathing 4D-CT acquisition techniques have been...
developed. An optical flow based temporal interpolation technique was introduced to reduce motion artifacts at the overlapping regions of images acquired at different couch positions during free breathing. A temporal nonlocal means strategy was employed by Tian et al. to regularize the reconstruction of 4D-CT data with undersampled projections. To reduce the number of phases that need to be scanned, a registration-based approach was applied to reconstruct intermediate phases based on a number of reference phases (e.g., inhale and exhale phases).

The main objective of this paper is to enhance the interslice resolution of lung 4D-CT data in order to help treatment planning for radiation therapy. For enhancing image resolution, super-resolution (SR) reconstruction is an effective postprocessing approach. Classical SR methods can be divided into two major categories: interpolation-based and model-based. Interpolation-based methods have two types: scene-based and object-based methods. Linear and spline-based interpolation methods are the commonly used scene-based approaches. The main advantage of scene-based methods is their simplicity. However, blurred edges and undesirable artifacts are often inevitable due to the use of simple averaging in these methods. The object-based methods are based on the anatomical structure similarity of consecutive slices to guide interpolation. Shape-based, morphology-based, optical-flow-based, and registration-based approaches are the typical object-based methods. However, object-based interpolation of lung CT data is often ineffective due to rapid structural changes across slices. Recently, research attention has been directed to model-based SR approach. The general assumption of model-based SR approaches is that the LR image is a degraded version of the SR image. The degradation can be represented using matrix representation such as $Z = DBM$, where $D$ is the down-sampling matrix, $B$ is the blur matrix, and $M$ is the transformation matrix, which can be used to characterize the effect of motion. More recent SR works aim to constrain the model with regularization terms, e.g., total variation, nonlocal means, etc. For instance, the Irani-Peleg backprojection model is employed to reconstruct SR MRI images via the estimation of the point spread function (PSF) associated with subject motion. The maximum a posteriori (MAP) algorithm is used to reconstruct high resolution (HR) synchronized PET data by using B-spline registration for transformation estimation. For the above approach, the degradation matrix $Z$ has direct influence on the SR reconstruction, and yet its estimation is nontrivial and often error-prone.

A more recent alternative is the use of learning based methods. The key idea of these methods is to utilize the relationship between the HR and LR images, learned via training images, to help recover details in the target LR images. For instance, Freeman et al. proposed an example-based super-resolution method which can preserve fine details in the recovered HR images, such as edges and plausible textures. Chang et al. used a neighboring embedding approach to reconstruct HR images. The HR image patches were recovered based on not only a nearest neighbor but also multiple nearest neighbors in the training set, in a way similar to locally linear embedding (LLE). Yang et al. employed a sparse representation method using an over-complete dictionary to reconstruct HR images. All these methods indicated that the state-of-the-art performance can be achieved by using various learning-based methods. However, a major limitation of these methods is that a significant amount of HR images is required for learning the relationship between the HR and LR images. In the case of lung 4D-CT SR reconstruction, the required HR dataset may not even be available.

In this paper, we propose a different method for enhancing the interslice resolution of 4D-CT lung data based on the concept that the complementary anatomical information may distribute throughout images acquired at different phases. Generally, 4D-CT captures 5–10 phases of the respiratory cycle, corresponding to different stages of motion of the lung. Therefore, information can be propagated between phases to recover structural details that are missing in one particular phase. It is worth noting that our approach can avoid artifacts introduced by averaging of dissimilar structures, does not require the estimation of the degradation matrix, and can obviate the need of high resolution data for training.

The core of our method is fuzzy patch matching. Specifically, inspired by recent advances in nonlocal image processing, e.g., for denoising, labeling, and segmentation, we propose a patch-based approach to help reconstruct a resolution enhanced 4D-CT data. The approach is divided into two steps. In the first step, the intermediate slices are reconstructed. At the beginning, a new patch distance measure is utilized for determining the matching patches. Then, a nonlocal strategy is employed to combine the matching patches for high-resolution reconstruction. For greater adaptation to anatomical structures of different scales, we utilize a quad-tree based algorithm to adaptively partition CT slices into patches of different sizes. In the second step, an iterative strategy is used to refine the results generated in the previous step. We will demonstrate both qualitatively and quantitatively that our method yields superior performance when compared with the conventional linear and cubic-spline interpolation methods.

The rest of this paper is organized as follows. In Sec. II, we will present the details of our method for enhancing the resolution of lung 4D-CT data. Experimental results on extensive evaluation of the proposed method in comparison with the conventional interpolation methods will be provided in Sec. III. We then give the discussion and conclusions in Sec. IV.

II. METHODS

II.A. Subjects

The data used in this paper are obtained from a publicly available dataset that was provided by the DIR-lab at the University of Texas M.D. Anderson Cancer Center (Houston, TX). The dataset consists of 10 cases of 4D-CT lung data, which were acquired as part of the standard planning process for the treatment of thoracic malignancies by employing a GE medical system (General Electric Discovery DT PET/CT...
For each case, the 4D-CT images covered the entire thorax and upper abdomen, and contained 10 phases, including the extreme inhale and exhale phases. For case 1 to case 5, the in-plane grid size is $256 \times 256$ and the in-plane voxel dimensions range from $(0.97 \times 0.97)$ to $(1.16 \times 1.16)$ mm$^2$. For case 6 to case 10, the in-plane grid size is $256 \times 512$ and the in-plane voxel dimensions are all $(0.97 \times 0.97)$ mm$^2$. For all cases, the image slice spacing along the superior–inferior direction is $2.5$ mm. Table I gives the detailed information regarding the dataset.

### II.C. Overview

The main goal of this work is to reconstruct intermediate slices of a particular phase using the information from other phases captured within the same 4D-CT data. Resolution enhancement is achieved by patch matching and composition with constraints from the neighboring slices. In contrast to the commonly used resolution enhancement approaches, the proposed method is simple, does not use any PSF estimation, and does not require training using a database of images.

Key components of the proposed approaches are as follows:

- A nonlocal means based approach to combine inter-phase matching patches for 4D-CT resolution enhancement.
- A quad-tree based strategy to partition the image into adaptive patches to deal with anatomies of different scales.

The main principles of our method are described in Sec. II.D. Sections II.E and II.F provide further details. A summary of the algorithm is provided in Sec. II.G.

### II.D. Principles of our method

Given an acquired 4D-CT image $I = \{I_i(s)\}_{i = 1, \ldots, N}$, where $N$ is the number of phases and $S$ is the total number of slices in each phase image $I_i$, the goal is to reconstruct an intermediate slice between two consecutive slices $I_i(s)$ and $I_i(s + 1)$ in image $I_i$. Our proposed solution is illustrated in Fig. 2. The larger dashed square denotes the intermediate slice to be reconstructed between slice $s$ and $s + 1$ in phase 1. The smaller box (denoted by $y$ in Fig. 2) denotes a patch in the intermediate slice that needs to be estimated. For each patch $y$ (smaller dashed box), we intend to employ a set of structurally similar patches, i.e., $y_1, y_2, \ldots, y_j$ (shown in boxes in Fig. 2), that are obtained from other phases (i.e., Phase 2, Phase 3, \ldots, Phase $N$) to reconstruct it. Two questions immediately arise: (1) How do we determine the suitable patches to form the patch set? (2) How do we combine matching patches effectively? Proposed solutions to these two problems will be discussed next.

**II.E. First step: Initial reconstruction of intermediate slices**

**II.E.1. Determining matching patches**

To determine patches that can be adopted to reconstruct $y$, we search for patches with similar structures in a limited spatial region in all other phases. Specifically, patch similarity with respect to a candidate patch $y'$ is simply evaluated based on the Euclidean distance:

$$D(y, y') = \|y - y'\|_2.$$

(1)

Based on the distance measurement, similar patches can be found. However, in our case, $y$ is not known beforehand. A possible approach is to first obtain an interpolated version of $y$ such as based on $y^a$ and $y^d$ (see Fig. 2), which are the

<table>
<thead>
<tr>
<th>Original dataset</th>
<th>Image dimensions</th>
<th>Voxel dimensions (mm)</th>
<th>Phase number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>$256 \times 256 \times 94$</td>
<td>$0.97 \times 0.97 \times 2.5$</td>
<td>10</td>
</tr>
<tr>
<td>Case 2</td>
<td>$256 \times 256 \times 112$</td>
<td>$1.16 \times 1.16 \times 2.5$</td>
<td>10</td>
</tr>
<tr>
<td>Case 3</td>
<td>$256 \times 256 \times 104$</td>
<td>$1.15 \times 1.15 \times 2.5$</td>
<td>10</td>
</tr>
<tr>
<td>Case 4</td>
<td>$256 \times 256 \times 99$</td>
<td>$1.13 \times 1.13 \times 2.5$</td>
<td>10</td>
</tr>
<tr>
<td>Case 5</td>
<td>$256 \times 256 \times 106$</td>
<td>$1.10 \times 1.10 \times 2.5$</td>
<td>10</td>
</tr>
<tr>
<td>Case 6</td>
<td>$512 \times 512 \times 128$</td>
<td>$0.97 \times 0.97 \times 2.5$</td>
<td>10</td>
</tr>
<tr>
<td>Case 7</td>
<td>$512 \times 512 \times 136$</td>
<td>$0.97 \times 0.97 \times 2.5$</td>
<td>10</td>
</tr>
<tr>
<td>Case 8</td>
<td>$512 \times 512 \times 128$</td>
<td>$0.97 \times 0.97 \times 2.5$</td>
<td>10</td>
</tr>
<tr>
<td>Case 9</td>
<td>$512 \times 512 \times 128$</td>
<td>$0.97 \times 0.97 \times 2.5$</td>
<td>10</td>
</tr>
<tr>
<td>Case 10</td>
<td>$512 \times 512 \times 120$</td>
<td>$0.97 \times 0.97 \times 2.5$</td>
<td>10</td>
</tr>
</tbody>
</table>

**Fig. 1.** Extracting the lung regions. Top: the original lung data. Bottom: the mask for the lung.
patches immediately superior and inferior to \( y \), using either linear interpolation or cubic-interpolation. However, as aforementioned, interpolation will result in undesirable artifacts in \( y \), which will subsequently influence patch matching significantly, as will be illustrated in Sec. III.

In this work, \( y^a \) and \( y^d \) are used as constraints for the reconstruction of \( y \). First, we expect that the reconstructed patch \( y \) should resemble \( y^a \) and \( y^d \). In this way, a joint distance between \( y' \) and \( y^a \) and between \( y' \) and \( y^d \) is defined:

\[
D(y^a, y^d, y') = \|y^a - y'\|^2_2 + \|y^d - y'\|^2_2. \tag{2}
\]

In addition, we require that the reconstructed patch is not biased toward any of the consecutive patches, i.e., \( \|y^a - y''\|^2_2 \) and \( \|y^d - y''\|^2_2 \) should be balanced:

\[
\frac{1}{\epsilon} \leq \frac{\|y^a - y''\|^2_2}{\|y^d - y''\|^2_2} \leq \epsilon, \tag{3}
\]

where \( \epsilon \) is a tolerance factor. For better characterization of structural patterns, we use both intensity and derived feature to represent each patch. Thus, Eqs. (2) and (3) can be combined as

\[
D(\tilde{y}^a, \tilde{y}^d, \tilde{y}') = \|\tilde{y}^a - \tilde{y}'\|^2_2 + \|\tilde{y}^d - \tilde{y}'\|^2_2 \quad \text{s.t.} \quad \frac{1}{\epsilon} \leq \frac{\|\tilde{y}^a - \tilde{y}'\|^2_2}{\|\tilde{y}^d - \tilde{y}'\|^2_2} \leq \epsilon, \tag{4}
\]

where \( \tilde{y}^a \), \( \tilde{y}^d \), and \( \tilde{y}' \) are the new patches, as defined by \( \tilde{y}^a = [\tilde{y}^a_{F_G}] \), \( \tilde{y}^d = [\tilde{y}^d_{F_G}] \), and \( \tilde{y}' = [\tilde{y}'_{F_G}] \).

Here, \( F \) is the feature operator. We use gradients as features in our approach. \( \lambda \) is a tuning parameter controlling the balance between the contributions of image intensity and image feature.

Based on Eq. (4), we can evaluate the effectiveness of each candidate patch \( y' \) by (a) calculating the joint distance between \( \tilde{y}^a \) and \( \tilde{y}' \) and between \( \tilde{y}^d \) and \( \tilde{y}' \); and (b) balancing \( \|\tilde{y}^a - \tilde{y}'\|^2_2 \) and \( \|\tilde{y}^d - \tilde{y}'\|^2_2 \) by calculating \( \frac{\|\tilde{y}^a - \tilde{y}'\|^2_2}{\|\tilde{y}^d - \tilde{y}'\|^2_2} \). The candidate patch with the lowest joint distance and also meeting the constraint \( \frac{1}{\epsilon} \leq \frac{\|\tilde{y}^a - \tilde{y}'\|^2_2}{\|\tilde{y}^d - \tilde{y}'\|^2_2} \leq \epsilon \) will be selected.

\[II.E.2. \text{Nonlocal approach}\]

Upon obtaining the fitting patches of \( y \), the next task is to determine a suitable reconstruction strategy to combine them together. In the recent work, Buades et al.\textsuperscript{24} showed that nonlocal means filtering gives the state-of-the-art performance in structure-preserving image denoising. The strategy has also been applied to brain image labeling,\textsuperscript{25} image registration,\textsuperscript{30} and MR image super-resolution.\textsuperscript{31} We employ nonlocal averaging for combining all matching patches that have been determined based on the distance measure as described in Sec. II.E.1. The nonlocal averaging is performed as below:

\[
y = \frac{\sum_{y' \in \Omega} w(\tilde{y}^a, \tilde{y}^d, \tilde{y}') y'}{\sum_{y' \in \Omega} w(\tilde{y}^a, \tilde{y}^d, \tilde{y}')}, \tag{5}
\]

where \( \Omega \) is the patch set composed of all matching patches. \( w \) is the weight that is associated with the distance \( D(\tilde{y}^a, \tilde{y}^d, \tilde{y}') \), which is computed as follows:

\[
w(\tilde{y}^a, \tilde{y}^d, \tilde{y}') = \exp \left( -\frac{D(\tilde{y}^a, \tilde{y}^d, \tilde{y}')}{2\sigma^2} \right), \tag{6}
\]

where \( \sigma \) controls the decay of the exponential function.

\[II.E.3. \text{Quad-tree based patch partition}\]

A commonly used approach to partition a slice is to divide it into identically sized patches. However, this approach fails to take into account the fact that anatomical structures are manifested in different scales. Consequently, we employ a quad-tree based strategy\textsuperscript{32} to partition the slice into structurally adaptive patches. Since the intermediate slice is not available beforehand, we partition the slices immediately superior and inferior to the intermediate slice. A standard top-down approach to construct the quad-tree is performed for each slice. Starting with the entire slice, we test each superior–inferior patch-pair (e.g., \( y^a \) and \( y^d \)) simultaneously to see if they meet a predefined homogeneity criterion. If the superior–inferior patch-pair meets the criterion, the division is halted. Otherwise, the patch-pair will be further divided into subpatches. This procedure is repeated iteratively until each patch-pair meets the stopping criterion. We employ a
FIG. 3. Quad-tree based adaptive patch partitioning of two consecutive slices.

simple intensity-based homogeneity criterion, where for the patch-pair the split stops if the patch-pair intensity variance is below a specified threshold. An example of the slice partitioning is shown in Fig. 3.

II.E.4. Determining the number of candidate patches

We use the above approach to reconstruct the initial intermediate slices. Figure 4 shows an example of a reconstructed slice. It is clear that our method achieves the best result. For example, the reconstructed slice shows the greatest similarity with the ground truth image and least artifact. More results will be reported in Sec. III.

To achieve the best reconstruction for the initial resolution-enhanced slices, we need to first determine the number of matching patches for $y$. We use four typical cases (case 1, case 3, case 5, and case 7 from the DIR-lab) to test the influence of this parameter.

Figure 5 shows the average peak signal-to-noise ratio (PSNR, a commonly used metric for evaluation of reconstruction error as detailed in Sec. III.F) for the reconstructed intermediate slices with respect to the use of different candidate fitting patch numbers. It can be seen that setting the candidate patch number to 6–9 provides the best reconstruction result in terms of PSNR. This is not surprising since too few candidate patches will provide too little information for reconstruction. On the other hand, too many candidate patches will also introduce image blurring. Therefore, we set the candidate patch number to 8 in the first step throughout all experiments presented in Sec. III. The second step of iterative refinement is introduced below.

II.F. Second step: Iterative refinement

We have mentioned previously that, in order to achieve better reconstruction results, the number of candidate patches should be set as much as 8. This will, however, blur the reconstructed intermediate slices due to the average approach. To overcome this problem, we iteratively decrease the candidate patch number, and further repeatedly update the reconstructed slices to reduce image blurring.

Let $Y'$ denote the initial reconstructed image and $y'$ denote a local patch of $Y'$. We search for patches with similar structures of $y'$ in a limited spatial region in all other phases. Patch similarity with respect to a candidate patch $y'$ is evaluated based on the Euclidean distance:

$$D(\bar{y}', \bar{y}) = \|\bar{y}' - \bar{y}\|^2_2.$$  \hspace{1cm} (7)

Note that this equation is similar to Eq. (4). The difference is that now we use the initial reconstructed patch $y'$ as constraint, instead of $\bar{y}^u$ and $\bar{y}^d$, and Eq. (3) is not required. $\bar{y}'$ and $\bar{y}'$ are also defined by $\bar{y}' = \lfloor \lambda F y' \rfloor$ and $\bar{y}' = \lfloor \lambda F y' \rfloor$. The nonlocal averaging is then performed as below:

$$y = \frac{\sum_{y' \in \Omega} w(\bar{y}', \bar{y}') y'}{\sum_{y' \in \Omega} w(\bar{y}', \bar{y}')},$$  \hspace{1cm} (8)

where the weight is calculated as

$$w(\bar{y}', \bar{y}') = \exp \left( -\frac{D(\bar{y}', \bar{y}')}{2\sigma^2} \right).$$  \hspace{1cm} (9)
We progressively reduce the candidate patch number, and then perform the above procedure iteratively to obtain the final result. Figure 6 shows an example of how the reconstructed image is progressively refined using the proposed iterative strategy. From left to right, results for decreasing candidate patch number are shown. Then, the enlarged views of the area marked by squares are shown in the second row for better visual comparison. It can be seen that the image becomes increasingly clearer in each step.

II.G. Summary of the algorithm

Now we summarize the proposed algorithm. Table II gives the detailed steps.

III. EXPERIMENTAL RESULTS

To evaluate the performance of our patch-based resolution enhancement method, we apply our method to the DIR-lab dataset (see Sec. II.A). We simulated the 4D-CT with 5 mm slice thickness by removing one of every two slices from the original 4D-CT with 2.5 mm slice thickness. Thus, we can quantitatively evaluate the effectiveness of our method in enhancing the resolution from 5 to 2.5 mm. We present both visual and quantitative results to demonstrate the performance of the proposed method, with the comparison to linear interpolation and cubic-spline interpolation. The algorithm was implemented using C++ language and all the following evaluations are performed using a computer with a 2.4 GHz processor and 3 GB of memory.

By comparing the ground truth image with the reconstructed intermediate slices, we first evaluate the influence of special strategies of our proposed method and the parameters.

III.A. The influence of using neighboring slices constraint

Figure 7 demonstrates the reconstruction results generated with our method. The top row shows the ground-truth slice together with its respective inferior and superior slices. The
bottom row shows the result based on linear interpolation, the reconstructed result based on Eq. (1) (where the linear interpolation result is used as reference for patch distance computation), and the result given by the proposed method [where patch distance is computed based on Eq. (4)]. It can be observed that our method provides better reconstruction result that resembles the ground truth [Fig. 7(b)] more closely. The patch matching due to the artifacts in the interpolated patch [see Fig. 7(d)] is avoided effectively.

Figure 8 shows the quantitative comparison of reconstructed local patch based on linear interpolation and the proposed method. It can be seen that both PSNR and structural similarity (SSIM, detailed information presented in Sec. III.F) increase significantly by using our method.

### III.B. The influence of using adaptively sized patch

Figure 9 shows the results based on the identically sized patches and our quad-tree based adaptively sized patches (described in Sec. II.E.3), respectively. It is clear that the vessels are more fidelity with the results given by the adaptively sized patch-based approach. This is because dividing the patches into difference sizes will catch more details with respect to vessel structures, and thus will help improve the reconstruction.

### III.C. The influence of parameter λ

The parameter λ balances the effects of image intensities and image features. Figure 10 shows the influence of the parameter λ for the computation of PSNR with the value changing from 0 to 0.8 for four cases (case 1, case 3, case 5, and case 7). Based on this experiment, it is clear that λ = 0.4 provides the best reconstruction result in terms of PSNR. It is reasonable that using the joint image intensity and feature information will find more fitting matches. However, too much feature will introduce artifacts and noise. We, thus, set λ to 0.4 throughout all experiments in Secs. III.E and III.F.

### III.D. The influence of parameter ε

The tolerance factor ε controls the bias of the similarity of the selected patches with the two consecutive patches (the immediate upper and lower patches). Similar with the analysis for the parameter λ, we also calculate the PSNR for the same four cases to test the influence of ε. Figure 11 shows the PSNR values with different ε ranging from 1.1 to 1.5. It appears that, for the low ε value, the higher PSNR is achieved. This is not surprising since strict constraint is helpful to reject irrelevant patches. Therefore, for all the experiments in Secs. III.E and III.F, we set ε to 1.1.

### III.E. Visual evaluation

In Fig. 12, we show typical reconstruction results given by different algorithms. The ground-truth images are shown in...
the left column. The results given by linear interpolation, cubic spline interpolation, and our method are shown in the second, third, and fourth column, respectively. The color-coded difference maps computed with respect to the ground truth are also shown to aid comparison. It is apparent that our proposed method outperforms the conventional interpolation methods, which often cause undesirable artifacts. Our method utilizes the information available from all other phase images to compensate the missing information and hence expectedly yields better results.

In Fig. 13, we further provide both coronal and sagittal views of the same 4D-CT case for comparison, with similar arrangement as Fig. 12. These results again demonstrate the best performance achieved by our proposed method. Regions in the circles are enlarged and are shown in the bottom of each panel, respectively. It can be more clearly observed that our method yields results with more accurate tumor shape and greater vessel continuity.

III.F. Quantitative comparison with other interpolation methods

To quantitatively evaluate the performance of our proposed method, we utilize two quality measures: PSNR and SSIM index. PSNR is widely used in signal and image processing to measure reconstruction error, which is defined as the ratio between the maximum intensity value of the ground truth image and the mean square error between the ground truth and reconstructed image. Logarithmic decibel scale is used to express PSNR due to the wide dynamic range of images. SSIM index is a metric that is more consistent with the human visual system. It is calculated based on the average, variance, and covariance of intensity values on various windows of the ground truth and the reconstructed images. Average PSNR and SSIM values for each 4D-CT case of different approaches are reported in Fig. 14. The results indicate that the initial reconstruction using our method yields significant increase in PSNR and SSIM over linear and cubic-spline interpolation (PSNR: typically 3–4 dB higher, SSIM: 3%–5% gain). PSNR and SSIM continue to increase with iterative refinement.

Further comparison is performed using a statistical relevance measure \( \tau \) defined as

\[
\tau = \begin{cases} 
+100 \times \left(1 - \frac{\text{MSD}_{\text{other}}}{\text{MSD}_{\text{proposed}}} \right), & \text{if } \text{MSD}_{\text{other}} > \text{MSD}_{\text{proposed}} \\
-100 \times \left(1 - \frac{\text{MSD}_{\text{other}}}{\text{MSD}_{\text{proposed}}} \right), & \text{if } \text{MSD}_{\text{other}} \leq \text{MSD}_{\text{proposed}}.
\end{cases}
\]
where MSD\text{proposed} is the mean square value difference between the ground-truth image and the image reconstructed by our method, i.e.,

\[
\text{MSD}_{\text{proposed}} = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{S} \sum_{s=1}^{S} \left\| I_i^{\text{true}}(s) - I_i^{\text{proposed}}(s) \right\|^2.
\]

This demonstrates that the initial reconstruction using our method yields significant increase in PSNR and SSIM over linear and cubic-spline interpolation (PSNR: typically 3–4 dB higher, SSIM: 3%–5% gain).
TABLE III. Statistical relevance values computed using MSD and NSD.

<table>
<thead>
<tr>
<th>Data</th>
<th>Resolution enhancement methods</th>
<th>MSD</th>
<th>NSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>Our/lin</td>
<td>34.68</td>
<td>39.41</td>
</tr>
<tr>
<td></td>
<td>Our/cubic</td>
<td>33.82</td>
<td>38.09</td>
</tr>
<tr>
<td>Case 2</td>
<td>Our/lin</td>
<td>32.36</td>
<td>35.09</td>
</tr>
<tr>
<td></td>
<td>Our/cubic</td>
<td>30.77</td>
<td>29.22</td>
</tr>
<tr>
<td>Case 3</td>
<td>Our/lin</td>
<td>40.38</td>
<td>47.90</td>
</tr>
<tr>
<td></td>
<td>Our/cubic</td>
<td>38.62</td>
<td>43.63</td>
</tr>
<tr>
<td>Case 4</td>
<td>Our/lin</td>
<td>36.09</td>
<td>43.01</td>
</tr>
<tr>
<td></td>
<td>Our/cubic</td>
<td>34.97</td>
<td>39.01</td>
</tr>
<tr>
<td>Case 5</td>
<td>Our/lin</td>
<td>31.14</td>
<td>36.32</td>
</tr>
<tr>
<td></td>
<td>Our/cubic</td>
<td>30.70</td>
<td>34.55</td>
</tr>
<tr>
<td>Case 6</td>
<td>Our/lin</td>
<td>24.24</td>
<td>23.32</td>
</tr>
<tr>
<td></td>
<td>Our/cubic</td>
<td>24.47</td>
<td>24.04</td>
</tr>
<tr>
<td>Case 7</td>
<td>Our/lin</td>
<td>36.53</td>
<td>37.52</td>
</tr>
<tr>
<td></td>
<td>Our/cubic</td>
<td>35.73</td>
<td>36.27</td>
</tr>
<tr>
<td>Case 8</td>
<td>Our/lin</td>
<td>30.32</td>
<td>29.43</td>
</tr>
<tr>
<td></td>
<td>Our/cubic</td>
<td>29.88</td>
<td>31.28</td>
</tr>
<tr>
<td>Case 9</td>
<td>Our/lin</td>
<td>33.71</td>
<td>32.54</td>
</tr>
<tr>
<td></td>
<td>Our/cubic</td>
<td>32.92</td>
<td>32.68</td>
</tr>
<tr>
<td>Case 10</td>
<td>Our/lin</td>
<td>28.54</td>
<td>21.45</td>
</tr>
<tr>
<td></td>
<td>Our/cubic</td>
<td>27.79</td>
<td>25.26</td>
</tr>
</tbody>
</table>

$T$ to $5\%$ of the maximum intensity value in the ground-truth image. For instance, $\text{NSD}_{\text{proposed}}$ is calculated as follows:

$$
\text{NSD}_{\text{proposed}} = \frac{1}{N} \sum_{i=1}^{N} \sum_{s=1}^{S} \sum_{v \in R} 1 \left( \frac{1}{I_{i}^{\text{true}}(s; v)} - \frac{1}{I_{i}^{\text{true}}(s; v)} \right)
$$

and

$$
\gamma = \begin{cases} 1, & \text{if } x > T \\ 0, & \text{otherwise}, \end{cases}
$$

where $v$ is pixel in a slice $I_{i}^{\text{true}}(s)$ and $R$ is the entire image domain.

Compared with linear interpolation (our/lin) and cubic-spline interpolation (our/cubic), the results, shown in Table III, again demonstrate that the proposed method yields superior performance.

IV. DISCUSSION AND CONCLUSION

In this paper, we have proposed a novel lung 4D-CT resolution enhancement algorithm. We take advantage of complementary image information that can be taken from images of different phases captured in 4D-CT to recover the missing structural information. Specifically, a patch-based nonlocal strategy, exploiting a new patch distance measure and an adaptively slice-partitioning strategy, is utilized to achieve the resolution enhancement reconstruction. The proposed method demonstrates consistent improvements over all conventional interpolation methods both qualitatively and quantitatively.

Image with sufficient resolution is a key to accurate subsequent processing such as image segmentation, three-dimensional reconstruction, and also clinical diagnosis and treatment planning. Factors such as acquisition time and radiation dose limit often hinder the acquisition of high-resolution images (such as ≤1 mm in longitudinal direction). Scene-based interpolation methods are simple and fast, but often cause artifacts and blur structures. Model-based interpolation algorithms\textsuperscript{10–13} reduce artifacts and preserve edges, but are limited by whether consecutive slices are sufficiently similar. Zhang et al.\textsuperscript{35} discussed to enhance 4D-CT resolution using registration-based interpolation approach. However, we note that, in lung 4D-CT data, structures differ significantly between consecutive slices. Hence, registration-based interpolation is unreliable since accurate structural alignment cannot be easily achieved.

The current work builds upon the work of Zhang et al.\textsuperscript{35} In their work, an adaptive dictionary of patches is collected for each voxel location, based on which a (NP-hard) sparse representation problem is solved to reconstruct for super-resolution reconstruction. The proposed method, however, uses a simple but effective nonlocal means based approach to combine matching patches for resolution enhancement. In addition, we employ an adaptive patch partition strategy to cater for different anatomical scales. Compared with Zhang et al.’s approach,\textsuperscript{35} our current method achieves better performance in a shorter amount of time. For example, to reconstruct an intermediate slice of size $256 \times 256$, the average time cost is 67.99 s using Zhang et al.’s approach, but only 33.97 s by our method. Zhang et al. employed several strategies for improving computation speed, such as using simple matching pursuit algorithm to solve the $L_1$ norm problem and building small dictionaries, instead of large dictionaries. Even so, our method is still faster (half the time requirement) and achieves better results. Table IV shows the PSNR and SSIM comparison of the proposed method and Zhang et al.’s approach.

As shown in Fig. 10, combining image intensity and image features in determining matching patches can significantly improve the final results. In this work, simple gradient features were used. More sophisticated features such as the integration features\textsuperscript{36} of Haar wavelets, histograms of gradients, and local binary patterns can potentially bring further

TABLE IV. PSNR and SSIM comparison of the proposed method and Zhang et al.’s approach.

<table>
<thead>
<tr>
<th>Data</th>
<th>Zhang et al. (Ref. 35)</th>
<th>Proposed</th>
<th>Zhang et al. (Ref. 35)</th>
<th>Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>33.01</td>
<td>33.77</td>
<td>0.9358</td>
<td>0.9469</td>
</tr>
<tr>
<td>Case 2</td>
<td>33.25</td>
<td>33.59</td>
<td>0.9278</td>
<td>0.9321</td>
</tr>
<tr>
<td>Case 3</td>
<td>34.55</td>
<td>35.53</td>
<td>0.9478</td>
<td>0.9552</td>
</tr>
<tr>
<td>Case 4</td>
<td>32.67</td>
<td>33.79</td>
<td>0.9486</td>
<td>0.9570</td>
</tr>
<tr>
<td>Case 5</td>
<td>32.52</td>
<td>33.20</td>
<td>0.9376</td>
<td>0.9480</td>
</tr>
<tr>
<td>Case 6</td>
<td>30.72</td>
<td>31.24</td>
<td>0.9151</td>
<td>0.9273</td>
</tr>
<tr>
<td>Case 7</td>
<td>30.78</td>
<td>31.54</td>
<td>0.9014</td>
<td>0.9163</td>
</tr>
<tr>
<td>Case 8</td>
<td>29.42</td>
<td>30.44</td>
<td>0.8475</td>
<td>0.8726</td>
</tr>
<tr>
<td>Case 9</td>
<td>32.22</td>
<td>32.78</td>
<td>0.9404</td>
<td>0.9463</td>
</tr>
<tr>
<td>Case 10</td>
<td>30.13</td>
<td>30.73</td>
<td>0.8999</td>
<td>0.9136</td>
</tr>
</tbody>
</table>
improvements. These features will be investigated in our future work.

The reconstruction accuracy of the current algorithm relies on matching with the immediate upper and lower patches. Future work includes the investigation of more complex matching strategies, such as group matching, for greater sensitivity to structural details and the further improvement of computational efficiency such as using multiresolution implementation. Instead of using Euclidean distance, our method can also be extended to employ geodesic or manifold distance.

Resolution enhancement of 4D-CT images will potentially benefit various clinical studies, e.g., CT-guided radiotherapy. The presented method tackles this problem by leveraging the correlated image information in different phases. Compared to previous studies that require training/registration/PSF function estimation, the proposed method provides a simple and elegant way to address this challenging problem. Our further work will be focused on improving the proposed method, as well as on the evaluation of feasibility of applying our method to facilitate accurate treatment planning and delivery of radiation therapy.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China under Grant Nos. 31271067 and 61072033, and the Major State Basic Research Development Program of China under Grant 2010CB732500. Q. Feng was supported by the National Major Scientific Instrument Equipment Development Program of China under Grant 2011YQ03011404.

REFERENCES


Medical Physics, Vol. 40, No. 5, May 2013

